arXiv:0704.0946v1 [astro-ph] 9 Apr 2007

EFFICIENT SIMULATIONS OF EARLY STRUCTURE FORMATION AND REIONIZATION

Andrei Mesinger & Steven Furlanetto
Yale Center for Astronomy and Astrophysics, Yale University, New Haven, CT 06520
Submitted to the ApJ

ABSTRACT

Detailed theoretical studies of the high-redshift universe, and especially reionization, are generally forced to rely on time-consuming N-body codes and/or approximate radiative transfer algorithms. We present a method to construct semi-numerical "simulations", which can efficiently generate realizations of halo distributions and ionization maps at high redshifts. Our procedure combines an excursion-set approach with first-order Lagrangian perturbation theory and operates directly on the linear density and velocity fields. As such, the achievable dynamic range with our algorithm surpasses the current practical limit of N-body codes by orders of magnitude. This is particularly significant in studies of reionization, where the dynamic range is the principal limiting factor because ionized regions reach scales of tens of comoving Mpc. We test our halo-finding and ionization-mapping algorithms *separately* against N-body simulations with radiative transfer and obtain excellent agreement. We compute the size distributions of ionized and neutral regions in our maps. We find even larger ionized bubbles than do purely analytic models at the same volume-weighted mean hydrogen neutral fraction, $\bar{x}_{HI}$, especially early in reionization. We also generate maps and power spectra of 21-cm brightness temperature fluctuations, which for the first time include corrections due to gas bulk velocities. We find that velocities widen the tails of the temperature distributions and increase small-scale power, though these effects quickly diminish as reionization progresses. We also include some preliminary results from a simulation run with the largest dynamic range to date: a 250 Mpc box that resolves halos with masses $M \geq 2.2 \times 10^8 M_\odot$. We show that accurately modeling the late stages of reionization, $\bar{x}_{HI} \lesssim 0.5$, requires such large scales. The speed and dynamic range provided by our semi-numerical approach will be extremely useful in the modeling of early structure formation and reionization.

Subject headings: cosmology: theory – early Universe – galaxies: formation – high-redshift – evolution

1. INTRODUCTION

Accurately modeling the formation of bound structures is invaluable for understanding any process in the early universe. Reionization, the epoch when radiation from early generations of astrophysical objects managed to ionize the intergalactic medium (IGM), is particularly sensitive to the distribution of collapsed structure. Current observations paint a complex picture of the reionization epoch (Mesinger & Haiman 2004; Wyithe & Loeb 2004; Fan et al. 2006; Mesinger & Haiman 2006; Malhotra & Rhoads 2004; Furlanetto et al. 2006c; Malhotra & Rhoads 2006; Page et al. 2006; Kashikawa et al. 2006; Totani et al. 2006). The next generation of instruments (*James Webb Space Telescope*; 21-cm instruments such as the Low Frequency Array and the Mileura Widefield Array Low-Frequency Demonstrator; CMB polarization measurements with *Planck*, etc.), could potentially shed light on this poorly understood milestone. Unfortunately, we still do not have accurate models of reionization with which to interpret these upcoming (*and* current) observations.

The main difficulty lies in the enormous dynamic range required. Ionized regions are expected to reach characteristic sizes of tens of comoving Mpc (Furlanetto et al. 2004c; Furlanetto & Oh 2005), which is over seven orders of magnitude in mass larger than the pertinent cooling mass, corresponding to gas with a temperature of $T \sim 10^4$ K (e.g. Efstathiou 1992; Thoul & Weinberg 1996; Gnedin 2000b; Shapiro et al. 1994). The required dynamic range is even larger if smaller "minihalos" be-

low this cooling threshold are important during reionization. Because of the steep mass dependence of halo abundances, halos with masses close to the cooling mass could dominate the photon budget. Hence modeling reionization requires simulation box sizes of hundreds of megaparsecs on a side, with extremely high resolution. Attempts to overcome these obstacles have generally followed the same fundamental and well-trod path (e.g. Gnedin 2000a; Razoumov et al. 2002; Ciardi et al. 2003; Sokasian et al. 2003; Iliev et al. 2006b; Zahn et al. 2007; Trac & Cen 2006): (1) N-body codes are run to generate halo distributions; (2) a simple prescription is used to relate the halo mass to an ionizing efficiency; (3) approximate methods (generally so-called ray-tracing algorithms) are used to model radiative transfer (RT) on large scales.

Even with modest halo resolution (Springel & Hernquist 2003) of tens of dark matter particles per halo, such schemes are computationally limited to box sizes of tens of megaparcecs, if they wish to resolve the likely cooling mass. McQuinn et al. (2006a) extended the mass resolution of their simulations by using a merger tree scheme to populate sub-grid scales with unresolved halos in a stochastic manner. Such hybrid schemes are useful for extending the dynamic range, but merger trees require a number of corrections to achieve consistent mass functions (see, e.g., Sheth & Pitman 1997; Benson et al. 2005, and Fig. 1 in McQuinn et al. 2006a) and to track individual halos with redshift. Moreover, although they are perfectly adequate for many purposes (including studying the

large-scale features of reionization), they prevent one from taking full advantage of the simulation.

Aside from dynamic range, the other main limiting factor in all of the above numerical approaches is speed. Even if the relevant scales can be resolved with N-body codes, such as may be the case in the early phase of reionization or with hybrid stochastic schemes. The codes themselves generally take days to run on large supercomputing clusters, with the approximate RT algorithms consuming a few additional days. The computational cost of each simulation makes it difficult to explore the full range of parameter space for reionization, which is particularly large because we know so little about high-redshift galaxies.

The computational cost becomes truly prohibitive if hydrodynamics is included: the largest such simulation of reionization performed to date spanned only $10h^{-1}$ Mpc (Sokasian et al. 2003). Including self-consistent descriptions of galaxy formation – even at the approximate level currently implemented in lower-redshift cosmological simulations (e.g., Springel & Hernquist 2003) – requires hydrodynamics, so N-body simulations of reionization are limited to semi-analytic prescriptions for star formation, feedback, etc. It is therefore worthwhile to explore even simpler schemes.

The purpose of this paper is to introduce approximate but efficient methods for generating halo distributions at high redshifts as well as for generating the associated ionization maps. We apply an excursion-set approach (e.g. Bond et al. 1991; Lacey & Cole 1993) to the filtering of a realization of the linear density field and then adjust halo locations with first-order perturbation theory. We can thus generate halo distributions at any given redshift, without explicitly including information from any higher redshifts. This scheme is an updated form of the "peak-patch" formalism developed and validated by Bond & Myers (1996a,b), although it was conceived and implemented completely independently. We then apply a similar technique to obtain the ionization field from the halo field. This part is similar to the schemes described in Zahn et al. (2005, 2007), except applied to our efficiently built halo distributions. As such, our methods allow us to make general predictions about non-linear processes, such as structure formation and reionization, without making use of time-guzzling cosmological simulations. The speed of our approach also allows us to explore a larger dynamic range than is possible with current cosmological simulations while preserving detailed spatial information (at least in a statistical sense), unlike purely analytic models.

This paper is organized as follows. In § 2, we introduce and test the components of our halo finding algorithm. In § 3, we introduce and test our HII bubble finding algorithm. In § 4, we use our semi-numerical scheme to generate maps and power spectra of expected 21-cm brightness temperature fluctuations throughout reionization. In § 5, we summarize our key findings and present our conclusions.

Unless stated otherwise, we quote all quantities in comoving units. We adopt the background cosmological parameters (Ω_Λ, Ω_M, Ω_b, n, σ_8, H_0) = (0.76, 0.24, 0.0407, 1, 0.76, 72 km s^{-1} Mpc^{-1}), consistent with the three–year results of the *WMAP* satellite (Spergel et al. 2006).

2. SEMI-NUMERICAL SIMULATIONS OF HALO PROPERTIES

In brief, our algorithm generates a linear density field and identifies halos within it. Because only linear evolution is required, the algorithm is fast and flexible. We generate 3D Monte-Carlo realizations of the linear density field on a box with sides of length $L = 100$ Mpc and $N = 1200^3$ grid cells. As such, we are able to take advantage of many pre-existing tools operating on the linear density field alone. Our method consists of the following principal steps:

1. creating the linear density and velocity fields
2. filtering halos from the linear density field using the excursion-set formalism
3. adjusting halo locations using their linear-order displacements

Step (1) only needs to be done once for each realization, since it is independent of redshift. As mentioned above, steps (2) and (3) need only be performed on redshifts of interest, i.e. since our output at redshift z is independent of any outputs at higher redshifts, there is no need for our code to "run down" to z, as is the case for N-body codes.

Our algorithm is an updated and simplified version of the "peak-patch" algorithm of Bond & Myers (1996a); we refer the interested reader there for more detailed explanations of some steps. A simpler version has also been used by Scannapieco et al. (2002) to study metal enrichment at high redshifts.

We perform our semi-numerical simulations on a single desktop Mac Pro with two dual-core 3.00 GHz Quad Xeon processors and 16 GB of RAM. For $N = 1200^3$, step (1) takes ~ 1 hour. For a given redshift in our range of interest, specifically for $z = 8.75$, steps (2) and (3) take ~ 2.5 hours. To achieve comparable halo mass resolution (including halos with $M \gtrsim 10^7 M_\odot$) with a minimum of ~ 500 particles per halo (Springel & Hernquist 2003), N-body codes would require a prohibitively large number of particles, $N \sim 10^{12}$! Below we describe in detail the components of our model.

2.1. *The Linear Density Field*

Our linear density field is generated in much the same way as it is for N-body codes. We briefly outline the procedure here.

The density field of the universe, $\delta(\mathbf{x}) \equiv \rho(\mathbf{x})/\bar{\rho} - 1$, in the linear regime[1] is well-represented as a Gaussian random field, whose statistical properties are fully defined by its power spectrum, $\sigma^2(\mathbf{k}) \equiv \langle |\delta(\mathbf{k})|^2 \rangle$. Here, $\delta(\mathbf{k})$ is the Fourier transform of $\delta(\mathbf{x})$, and the standard assumption of isotropy implies $\sigma^2(\mathbf{k}) = \sigma^2(k)$ while homogeneity implies that there are no density fluctuations with wavelengths larger than the box size $L = V^{1/3}$. We use the following, standard (e.g. Bagla & Padmanabhan 1997; Sirko 2005) Fourier transform conventions:

$$\delta(\mathbf{k}) = \frac{V}{N} \sum \delta(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} , \qquad (1)$$

---

[1] In linear theory, density perturbations evolve in redshift as $\delta(z) = \delta(0)D(z)$, where $D(z)$ is the linear growth factor normalized so that $D(0) = 1$ (e.g., Liddle et al. 1996). Unless the redshift dependence is noted explicitly, from this point forward we will work with quantities linearly-extrapolated to $z = 0$.

with the inverse transform being

$$\delta(\mathbf{x}) = \frac{1}{V} \sum \delta(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} . \qquad (2)$$

The discrete simulation box only permits a finite set of wavenumbers: $\mathbf{k} = \Delta k(i, j, k)$, where $\Delta k = 2\pi/L$ and i, j, k are integers in the range ($-\sqrt[3]{N}/2, \sqrt[3]{N}/2$]. For each independent wavenumber,[2] we assign

$$\delta(\mathbf{k}) = \sqrt{\frac{\sigma^2(k)}{2}}(a_{\mathbf{k}} + ib_{\mathbf{k}}) , \qquad (3)$$

where $a_{\mathbf{k}}$ and $b_{\mathbf{k}}$ are drawn from a zero-mean Gaussian distribution with unit variance. We use the power spectrum from Eisenstein & Hu (1999). Then the real-space density field, $\delta(\mathbf{x})$, is obtained by performing an inverse Fourier transform on $\delta(\mathbf{k})$.

2.2. *The Linear Velocity Field*

We construct a linear velocity field corresponding to our linear density field using the standard Zel'Dovich approximation (c.f. Zel'Dovich 1970; Efstathiou et al. 1985; Sirko 2005):

$$\mathbf{x_1} = \mathbf{x} + \psi(\mathbf{x}) , \qquad (4)$$

$$\mathbf{v} \equiv \dot{\mathbf{x}}_1 = \dot{\psi}(\mathbf{x}) , \qquad (5)$$

$$\delta(\mathbf{x}) = -\nabla\cdot[(1 + \delta(\mathbf{x}))\psi(\mathbf{x})]$$
$$\approx -\nabla \cdot \psi(\mathbf{x}) , \qquad (6)$$

where $\mathbf{x}$ and $\mathbf{x_1}$ denote initial (Lagrangian) and updated (Eulerian) coordinates, respectively, $\psi(\mathbf{x})$ is the displacement vector, and the last equation follows from the continuity criterion, with the final approximation using linearity, $\delta(\mathbf{x}) \ll 1$. We note again that all units are comoving, unless stated otherwise. From the above, one can relate the velocity mode in our simulation at redshift z to the linear density field:

$$\mathbf{v}(\mathbf{k}, z) = \frac{i\mathbf{k}}{k^2}\dot{D}(z)\delta(\mathbf{k}) , \qquad (7)$$

where for computational convenience differentiation is performed in k-space.

Another convenient property of this first-order Zel'Dovich approximation is that the velocity field can be decomposed into purely spatial, $\mathbf{v}_x(\mathbf{x})$, and purely temporal, $v_z(z)$, components:

$$\mathbf{v}(\mathbf{x}, z) = v_z(z)\mathbf{v}_x(\mathbf{x}) , \qquad (8)$$

where $v_z(z) = \dot{D}(z)$ and $\mathbf{v}_x(\mathbf{x})$ is the inverse Fourier transform of $i\mathbf{k}\delta(\mathbf{k})/k^2$. This is computationally convenient, as we only need to compute the $\mathbf{v}_x(\mathbf{x})$ field once in order to be able to scale it for all redshifts, and it also allows us to write a simple, exact expression for the *integrated* linear displacement field, $\mathbf{\Psi}$. When eq. (8) is integrated from some large initial z_0 [$D(z_0) \ll D(z)$], the total displacement is just

$$\mathbf{\Psi}(\mathbf{x}, z) = [D(z) - D(z_0)]\mathbf{v}_x(\mathbf{x})$$
$$\approx D(z)\mathbf{v}_x(\mathbf{x}) \qquad (9)$$

---

[2] Since $\delta(\mathbf{x})$ is real-valued, only half of the k-modes defined above are independent. The other half are determined by the usual Hermitian constraints for real-valued functions (see for example Hockney & Eastwood 1988; Bagla & Padmanabhan 1997).

We make use of this displacement field to adjust the halo locations obtained by our filtering procedure (see § 2.4), as well as to adjust the linear density field for our 21-cm temperature maps (see § 4).

In principle, one could obtain non-linear velocities by mapping the linear overdensity to a corresponding non-linear overdensity obtained from a spherical collapse model (Mo & White 1996), and then taking the time derivative of the non-linear overdensity. However, due to the large spread in the dynamical times of the non-linear density field, accurately capturing the time evolution is non-trivial. Furthermore, although the non-linear density field implicitly captures the velocities of collapsing gas, mapping each pixel's linear density to its non-linear counterpart independently of other nearby pixels does not properly preserve correlations on larger scales. Hence, we choose to use the linear density field directly in estimating velocities. For the purposes of studying the ionization field, we are further justified in this procedure because our final ionization maps are smoothed on large scales, on which most pixels are still in the linear regime at the high redshifts of interests. It is possible to include higher-order contributions to the Zel'Dovich approximation where necessary (e.g., Scoccimarro & Sheth 2002).

2.3. *Halo Filtering*

In standard Press-Schechter theory (PS; see e.g., Press & Schechter 1974; Bond et al. 1991; Lacey & Cole 1993), the halo mass function can be written as

$$\frac{\partial n(> M, z)}{\partial M} = -\sqrt{\frac{2}{\pi}}\frac{\bar{\rho}}{M}\frac{\delta_c(z)}{\sigma^2(M)}\frac{\partial\sigma(M)}{\partial M}\exp\left[-\frac{\delta_c^2(z)}{2\sigma^2(M)}\right], \qquad (10)$$

where $n(> M, z)$ is the mean number density of halos with total mass greater than M, $\bar{\rho} = \Omega_{\mathrm{M}}\rho_{\mathrm{crit}}$ is the mean background matter density, $\delta_c(z) \sim 1.68/D(z)$ is the scale-free critical over–density evaluated in the case of spherically symmetric collapse (Peebles 1980), and

$$\sigma^2(M) = \frac{1}{V}\int_0^\infty \frac{k^2 dk}{2\pi^2}\sigma^2(k)W^2(k, M) , \qquad (11)$$

is the squared *r.m.s.* fluctuation in the mass enclosed within a region described by the filter function, $W(k, M)$, normalized to integrate to unity.

Although the PS mass function in eq. (10) is in fair agreement with simulations, especially for halos near the characteristic mass, at low redshifts it underestimates the number of high–mass halos and overestimates the number of low–mass halos when compared with large numerical simulations (e.g. Jenkins et al. 2001). A modified expression shown to fit low-redshift simulation results more accurately (to within $\sim 10\%$) was obtained by Sheth & Tormen (1999):

$$\frac{\partial n(> M, z)}{\partial M} = -\frac{\bar{\rho}}{M}\frac{\partial[\ln \sigma(M)]}{\partial M}\sqrt{\frac{2}{\pi}}A\left(1 + \frac{1}{\hat{\nu}^{2p}}\right)\hat{\nu}e^{-\hat{\nu}^2/2}, \qquad (12)$$

where $\hat{\nu} \equiv \sqrt{a}\delta_c(z)/\sigma(M)$, and a, p, and A are fitting parameters. Sheth et al. (2001) derive this form of the mass function by including shear and ellipticity in modeling non–linear collapse, effectively changing the scale-free critical over–density $\delta_c(z)$, into a function of filter scale,

$$\delta_c(M,z) = \sqrt{a}\,\delta_c(z)\left[\,1 + b\left(\frac{\sigma^2(M)}{a\delta_c^2(z)}\right)^c\,\right]. \qquad (13)$$

Here b and c are additional fitting parameters (a is the same as in eq. 12). For the constants above, we adopt the recent values obtained by Jenkins et al. (2001), who studied a large range in redshift and mass: $a = 0.73$, $A = 0.353$, $p = 0.175$, $b = 0.34$, $c = 0.81$. We note, however, that the situation at high redshifts is less clear: studies disagree on the relative accuracy of the Press & Schechter (1974) and Jenkins et al. (2001) forms (Reed et al. 2003; Iliev et al. 2006b; Zahn et al. 2007). Our algorithm can be trivially modified to accommodate other choices for the mass function; fortunately, for the purposes of the ionization maps (see §3), the choice of mass function makes very little difference because all have a similar dependence on the local density (Furlanetto et al. 2006a).

The mass functions in equations (10) and (12) can be obtained by the standard excursion set random walk procedure. The approach is to smooth the density field around a point, $\mathbf{x}$, on successively smaller scales starting with $M \to \infty$ [where $\sigma^2(M) \to 0$] and to identify the point as belonging to the halo with the largest M such that $\delta(\mathbf{x}, M) > \delta_c(M,z)$. If $W^2(k,M)$ is chosen to have a sharp cut-off, this procedure amounts to a random walk of $\delta(\mathbf{x}, M)$ along the mass axis, since the *change* in $\delta(\mathbf{x}, M)$ as the scale is shrunk is independent of $\delta(\mathbf{x}, M)$ for a top-hat filter in k-space (see eq. 11).

We perform this procedure on our realization of the linear density field by filtering the field using a real-space top-hat filter[3], starting on scales comparable to the box size and going down to grid cell scales, in logarithmic steps of width $\Delta M/M = 1.2$.[4] At each filter scale, we use the scale-dependent barrier in eq. (13) to mark a collapsed halo if $\delta(\mathbf{x}, M) > \delta_c(M,z)$. Filter scales large enough that collapsed structure is extremely unlikely, $\delta_c(M,z) > 7\sigma(M)$, are skipped (Mesinger et al. 2005). Since this procedure treats each cell as the center of a spherical filter, neighboring pixels are not properly placed in the same halo. Because of this, we discount halos which overlap with previously marked halos.

As mentioned above, this algorithm is similar to the "peak-patch" approach first introduced by Bond & Myers (1996a). The primary differences are:

---

[3] There is a slight swindle in the current application of this formalism. The filter function is assumed to be a top-hat in k-*space* in order to facilitate the analytic random walk approach described above. However, when the power spectrum is normalized to observations [i.e. $\sigma(R = 8h^{-1}\mathrm{Mpc}) = \sigma_8$), the filter that is used to define the mass M corresponding to R is a top-hat filter in *real space*. Nevertheless, it has been shown that the mass function is not very sensitive to this filter choice (Bond et al. 1991).

[4] We note that Mesinger et al. (2005) required a much smaller step size at these redshifts, $\Delta M/M \sim 0.1$, in order to produce accurate mass functions using 1D Monte-Carlo random walks. However, here we find that we can reproduce accurate mass functions with a larger step size, since in our 3D realization of the density field, "overstepping" $\delta_c(M,z)$ due to a large filter step size can be compensated with a small offset in the filter center, i.e. by centering the filter in a neighboring cell. This is the case since overstepping $\delta_c(M,z)$ means that some dense matter between the two filter scales was "missed". In a 1D Monte-Carlo random walk this matter is unrecoverable; however, in a 3D realization of the density field, the missed matter will be picked up by a filter centered on a neighboring cell.

(1) we use the Jenkins et al. (2001) barrier to identify halos (rather than calculating the strain tensor to account for ellipsoidal collapse), (2) we do not separately identify peaks in the density field (this step is not required given modern computing power), and (3) we use the "full exclusion" criterion for preventing halo overlap. Bond & Myers (1996b) found that a "binary exclusion" method in which pairs of overlapping halos are compared and eliminated was somewhat more accurate. However, at the high redshifts of interest to us, halo overlap is rare, and we are primarily interested in the large-scale properties of the halo field, which are relatively insensitive to the details of the overlap criterion.

We also note that our halo finder is similar in spirit to the PTHalos algorithm introduced by Scoccimarro & Sheth (2002) to generate mock galaxy surveys at low redshifts. There are two key differences. First, at present we use only first-order perturbation theory to displace the particles.[5] This limits us to higher redshifts, where velocities are smaller. However, our algorithm does not require particles in order to resolve halos and hence can accommodate a considerably larger dynamic range than PTHalos.

Mass functions resulting from this procedure are shown as points in Figure 1, with error bars indicating 1-σ Poisson uncertainties and bin widths spanning our mass filter steps. Dotted red curves denote PS mass functions generated by eq. (10); short–dashed blue curves denote extended PS conditional mass functions generated by eq. (10) but also taking into account the absence of density modes longer than the box size; long–dashed green curves denote mass functions generated using the Sheth-Tormen correction in eq. (12). The upper (lower) set of curves and points correspond to redshifts of $z{=}6.5$ ($z{=}10$). The dotted and short–dashed curves overlap at these redshifts due to our large box size ($L = 100$ Mpc), so we are immune to the finite box effects pointed out by Barkana & Loeb (2004).

Fig. 1 shows that we obtain accurate mass functions for $M \gtrsim 10^8 M_\odot$. Our procedure seems to underpredict the abundance of halos with masses approaching the cell size, $M_{\rm cell} \sim 10^7 M_\odot$. However, as the Jeans mass corresponding to a gas temperature of $\sim 10^4$ K is $M_J(z \sim 8) \sim 10^8 M_\odot$, in subsequent calculations, we only use halos with masses greater than $M_{\rm min} = 10^8 M_\odot$. Using this $M_{\rm min}$, we match the collapse fraction obtained by integrating eq. (12) to better than $\sim 10\%$.

This mass cutoff corresponds to the minimum temperature required for efficient atomic hydrogen cooling and would be the pertinent mass scale if: (1) the H_2 cooling channel is suppressed, e.g. due to a pervasive Lyman-Werner (LW) background, and if (2) photoionization feedback is ineffective at suppressing gas cooling and collapse onto higher mass halos. While feedback at high redshifts remains poorly-constrained, both of these assumptions seem reasonable during the middle stages of reionization on which we focus. A dissociating LW background is likely to have established

---

[5] We note that a similar scheme to ours has been independently created by O. Zahn (private communication). This scheme uses a simple Press-Schechter barrier but adjusts halo locations following second-order Lagrangian perturbation theory. However, he has found that the second-order corrections make very little difference to the map.
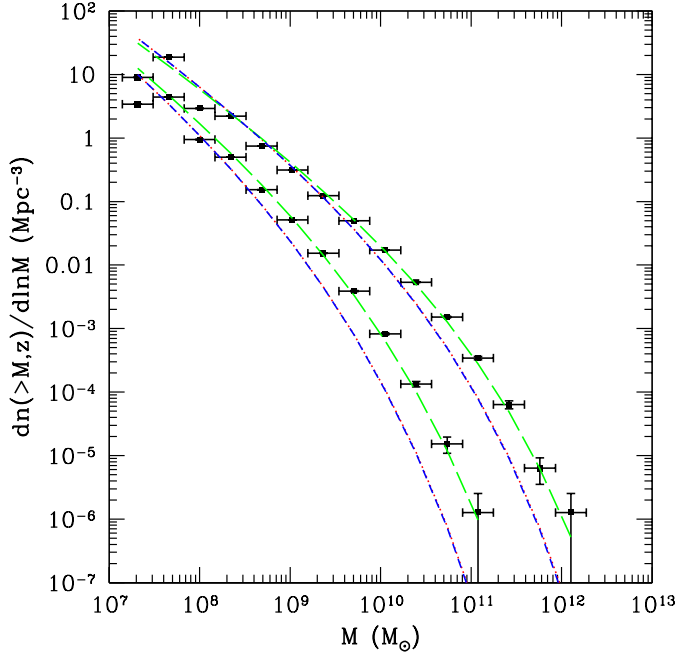
FIG. 1.— Mass functions generated from our halo filtering procedure discussed in §2.3 are shown as points. Dotted red curves denote PS mass functions generated by eq. (10); short–dashed blue curves denote extended PS conditional mass functions generated by eq. (10) but also taking into account the absence of density modes longer than the box size; long–dashed green curves denote mass functions generated using the Sheth-Tormen correction in eq. (12). The upper (lower) set of curves and points correspond to redshifts of z=6.5 (z=10).
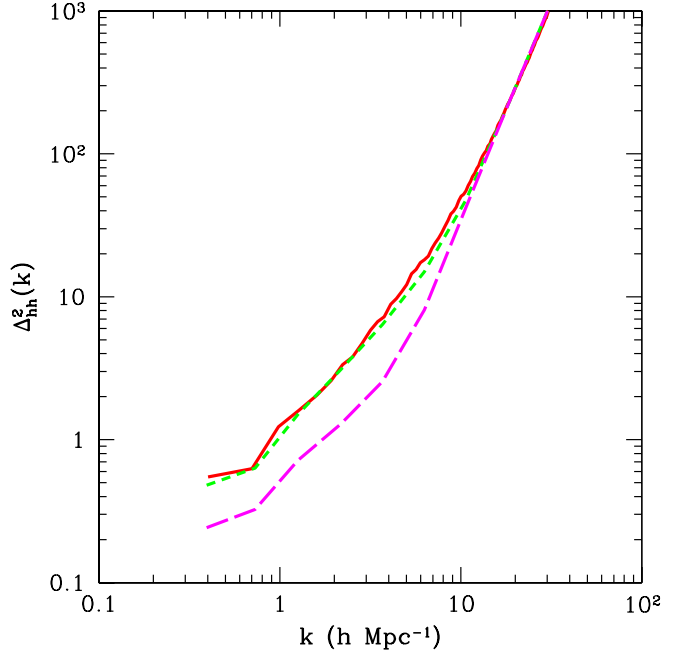

FIG. 2.— Halo power spectra at $z = 8.7$, with $L = 20 \ h^{-1}$Mpc and cosmological parameters taken from McQuinn et al. (2006a). The solid red curve is the halo power spectrum from an N-body simulation obtained from McQuinn et al. (2006a) (c.f. the bottom panel of their Fig. 2). The short-dashed green and the long-dashed violet curves are obtained from our filtering procedure with and without the halo location adjustments, respectively.

itself well before the universe is significantly ionized (Haiman et al. 1997). Model-dependent empirical evidence supporting the suppression of star formation in smaller mass halos, $M \lesssim M_{\min}$, can also be gleaned from *WMAP* data (Haiman & Bryan 2006). Furthermore, although early work suggested that an ionizing background could partially suppress star formation in halos with virial temperatures of $T_{\rm vir} \lesssim 3.6 \times 10^5$ K ($M \lesssim 2 \times 10^9 M_\odot$) (Thoul & Weinberg 1996), more recent studies (Kitayama & Ikeuchi 2000; Dijkstra et al. 2004) find that at high redshifts ($z \gtrsim 3$), self-shielding and the increased cooling efficiency could be strong countering effects for halos with virial temperatures $T_{\rm vir} > 10^4$ K. We postpone a more detailed analysis of the reionization footprint left by photo-ionization feedback to a future work.

2.4. *Adjusting Halo Locations*

Once the halo field is obtained, we use the displacement field obtained through eq. (9) to adjust the halo locations at each redshift. This corrects for the enhanced halo bias in Eulerian space with respect to our filtering, which is done in Lagrangian space (i.e. using the initial locations at large z). For computational convenience, we smooth the 1200^3 velocity field onto a coarser-grained 200^3 grid before adjusting halo locations. The choice of resolution, where each cell is $(100 \ {\rm Mpc})/200 = 0.5$ Mpc on a side, is somewhat arbitrary here, and we have verified that our halo and 21-cm power spectra are unaffected

by this choice. We also note that in linear theory, the mean velocity dispersion inside a $(0.5 \ {\rm Mpc})^3$ sphere with mean density at $z = 10$ is a factor of ~ 10 lower than the *r.m.s.* bulk velocity of such regions, so smoothing over smaller scale velocities appears reasonable. Furthermore, we keep in mind that our "endproducts" in this work are ionization and 21-cm temperature fluctuation maps, for which such "low-resolution" is more than adequate (compare, e.g., to N-body simulations of reionization, which typically have similar cell sizes for the radiative transfer component).

In Figure 2, we plot the halo power spectrum, defined as $\Delta_{\rm hh}(k, z) = k^3/(2\pi^2 V) \ \langle |\delta_{\rm hh}(\mathbf{k}, z)|^2 \rangle_k$, where $\delta_{\rm hh}(\mathbf{x}, z) \equiv M_{\rm coll}(\mathbf{x}, z)/\langle M_{\rm coll}(z) \rangle - 1$ is the collapsed mass field.[6] The solid red curve is the halo power spectrum from a 20 h^{-1} Mpc N-body simulation at $z = 8.7$ obtained from McQuinn et al. (2006a) (c.f. the bottom panel of their Figure 2). The short-dashed green and the long-dashed violet curves are obtained from our filtering procedure (matching the assumed cosmology) with and without the halo location adjustments, respectively. We note that ignoring the cumulative motions of halos results in an underestimate of the power of long-wavelength modes of the halo field by a factor of ~ 2 in this case. The average Eulerian bias of these halos is ~ 2, about half of which comes from the correction from Lagrangian to Eulerian coordinates.

After the halo locations are adjusted according to linear theory, our halo power spectrum agrees almost per-

---

[6] We use the collapsed mass field, rather than the individual galaxies, because we calculate the power from the smoothed cells.

fectly with the simulation. By design our procedure includes Poisson fluctuations in the halo number counts, which dominate the power spectrum at $k \gtrsim 5 \, h/\mathrm{Mpc}$ and are lost in purely analytic estimates (McQuinn et al. 2006a). We also note that both the halo mass functions and power spectra are statistical tests and hence the agreement shown here does not imply that our halo field has a one-to-one mapping with an N-body halo field sourced by identical initial conditions. Indeed, Gelb & Bertschinger (1994) showed that those particles located nearest initial linear density peaks are not necessarily incorporated into massive galaxies. The "peak particle" algorithm is less robust than our smoothing technique, but we still do not expect to recover halo masses or locations precisely. We plan on doing a "one-on-one" comparison between halo fields obtained from our halo finder to those obtained from N-body codes in a future work. However, it is certainly encouraging that the very similar "peak-patch" group finding formalism of Bond & Myers (1996a) did very well when compared "one-on-one" to N-body codes at large mass scales (Bond & Myers 1996b).

In Figure 3 we show slices through the halo field from our simulation box at $z = 8.25$, generated by the above procedure, again with (*right panel*) and without (*left panel*) the halo location adjustments. In the figure, the halo field is mapped to a lower resolution 400^3 grid for viewing purposes. Each slice is 100 Mpc on a side and 0.25 Mpc deep. Collapsed halos are shown in blue. Visually, it is obvious that peculiar motions increase halo clustering.

3. GENERATING THE IONIZATION FIELD

Once the halo field is generated as described above, we can perform a similar filtering procedure (also using the excursion-set formalism) to obtain the ionization field (similar methods have been discussed by Zahn et al. 2005, 2007). The time required for this final step is a function of $\bar{x}_{\mathrm{HI}}$, with large $\bar{x}_{\mathrm{HI}}$ requiring less time than small $\bar{x}_{\mathrm{HI}}$. Specifically, at $\bar{x}_{\mathrm{HI}} \sim 0.5$ this step takes ~ 15 minutes to generate a 200^3 ionization box on our workstation.

There are two main differences between the halo filtering and the HII bubble filtering procedures: (1) HII bubbles are allowed to overlap, and (2) the excursion set barrier (the criterion for ionization) becomes, as per Furlanetto et al. (2004a):

$$f_{\mathrm{coll}}(\mathbf{x_1}, M, z) \geq \zeta^{-1} \, , \tag{14}$$

where ζ is some efficiency parameter and $f_{\mathrm{coll}}(\mathbf{x_1}, M, z)$ is the fraction of mass residing in collapsed halos inside a sphere of mass $M = 4/3\pi R^3 \bar{\rho}[1 + \langle \delta_{\mathrm{nl}}(\mathbf{x_1}, z)\rangle_R]$, with mean physical overdensity $\langle \delta_{\mathrm{nl}}(\mathbf{x_1}, z)\rangle_R$, centered on Eulerian coordinate $\mathbf{x_1}$, at redshift z.

Equation (14) is only an approximate model and makes several simplifying assumptions about reionization. In particular, it assumes a constant ionizing efficiency per halo and ignores spatially-dependent recombinations and radiative feedback effects. It can easily be modified to include these effects (e.g., Furlanetto et al. 2004b, 2006a; Furlanetto & Oh 2005), and we plan to do so in future work. Here we present the simplest case in order to best match current RT numerical simulations.

This prescription models the ionization field as a two-phase medium, containing fully-ionized regions (which we refer to as HII bubbles) and fully-neutral regions. This is obviously much less information than can be gleaned from a full RT simulation, which precisely tracks the ionized fraction. However, HII bubbles are typically highly-ionized during reionization, and for many purposes (such as for 21 cm maps) this two-phase approximation is perfectly adequate.

In order to "find" the HII bubbles at each redshift we smooth the halo field onto a 200^3 grid. Then we filter the halo field using a real-space top-hat filter, starting on scales comparable to the box size and decreasing to grid cell scales in logarithmic steps of width $\Delta M/M = 0.33$. At each filter scale, we use the criterion in eq. (14) to check whether the region is ionized. If so, we flag all pixels inside that region as ionized. We do this for all pixels and scales, regardless of whether the resulting bubble would overlap with other bubbles. Note, therefore, that the nominal ionizing efficiency ζ that we use as an input parameter does not equal $(1 - \bar{x}_{\mathrm{HI}})/f_{\mathrm{coll}}$. They typically differ by $\lesssim 30\%$, with $\zeta f_{\mathrm{coll}} < 1 - \bar{x}_{\mathrm{HI}}$ early in reionization and $\zeta f_{\mathrm{coll}} > 1 - \bar{x}_{\mathrm{HI}}$ late in reionization). Unfortunately, we thus cannot use our algorithm to self-consistently predict the time evolution of the ionized fraction (rather, that must be prescribed from some other model). Of course, the same is true for N-body simulations, because the evolution of the ionized fraction depends on the evolving ionization efficiency of galaxies and cannot be self-consistently included in any present-day simulation.

In order to obtain the density field used in eq. (14), $\delta_{\mathrm{nl}}(\mathbf{x_1}, z)$, we use the Zel'Dovich approximation on our linear density field, $\delta(\mathbf{x})$, in much the same manner as we did to adjust our halo field in § 2.4. Starting at some arbitrarily large initial redshift (we use $z_0 = 50$), we discretize our high-resolution 1200^3 field into "particles" whose mass equals that in each grid cell. We then use the displacement field (eq. 9) to move the particles to new locations at each redshift. This resulting mass field is then smoothed onto our lower resolution 200^3 box to obtain $\delta_{\mathrm{nl}}(\mathbf{x_1}, z)$. We then recalculate the velocity field (§ 2.2) using the new densities.

Zahn et al. (2007) showed that a very similar HII bubble filtering procedure performed on an N-body halo field was able to reproduce the ionization topology obtained through a ray-tracing RT algorithm fairly well. Their algorithm differs from ours in two ways. First, they used a slightly different barrier definition; however, this difference has only a small impact on the ionization topology.[7] More importantly, for each filter scale at each pixel, Zahn et al. (2007) flag only the center pixel as ionized if the barrier is crossed, whereas we flag the entire filtered sphere.

In order to test our bubble filtering algorithm, we execute it on the same N-body halo field at $z = 6.89$ as was used to generate the bottom panels of Fig. 3 in Zahn et al. (2007). We compare analogous ionization maps created using various algorithms in Figure 4. All

---

[7] Specifically, in order to match the physics of their simulations better, they required $\int dt \, f_{\mathrm{coll}} > \zeta^{-1}$. However, the density modulation ends up nearly identical to our model, so the topology is almost unchanged.
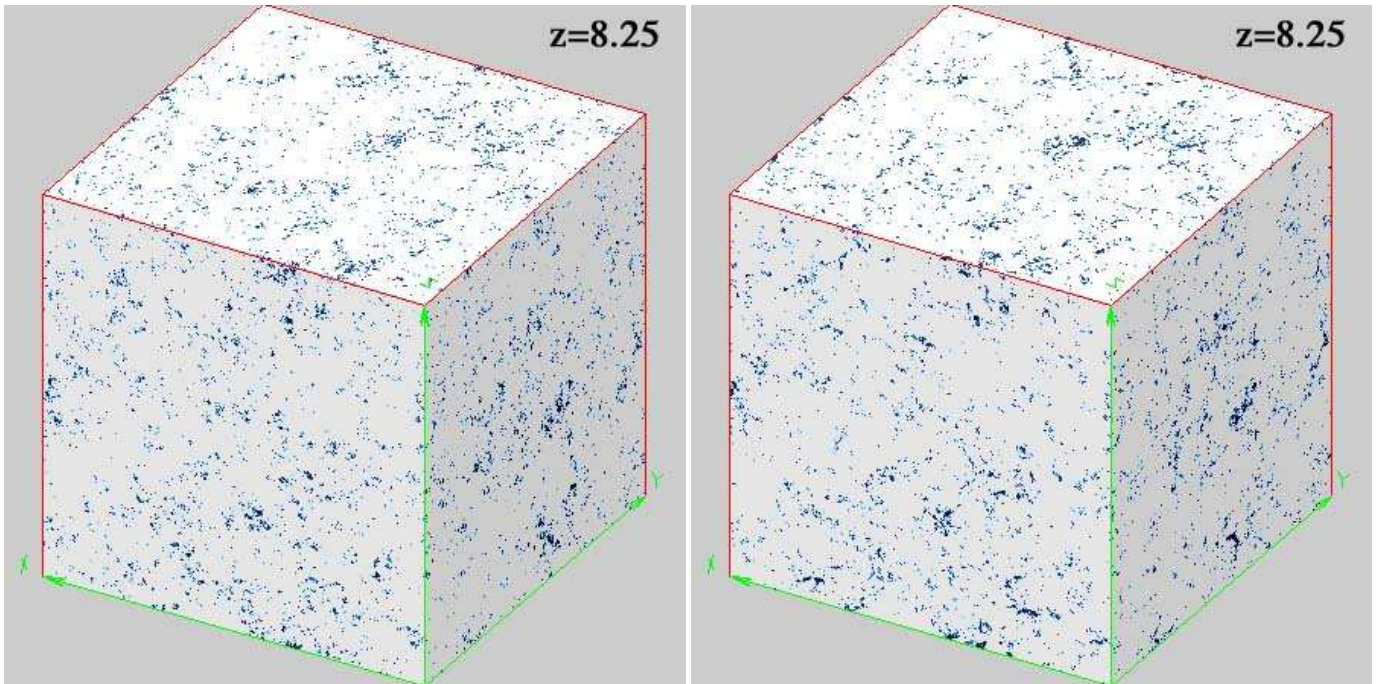
FIG. 3.— Slices through the halo field from our simulation box at $z = 8.25$. The halo field is generated on a 1200^3 grid and then mapped to a 400^3 grid for viewing purposes. Each slice is 100 Mpc on a side and 0.25 Mpc deep. Collapsed halos are shown in blue. The left panel shows the halo field directly filtered in Lagrangian space; the right panel maps the field to Eulerian space according to linear theory (see § 2.4 and eq. 9). The right panel corresponds to the bottom-left ($\bar{x}_{HI} = 0.53$) ionization field in Figure 5.

slices are 93.7 Mpc on a side and 0.37 Mpc deep, with ζ adjusted so that the mean neutral fraction in the box is $\bar{x}_{HI} = 0.49$. Ionized regions are shown as white. The leftmost and right-most panels are taken from Zahn et al. (2007). The left-most panel was created by performing their bubble filtering procedure directly on the linear density field (without explicitly identifying halos). The second panel was created by performing their bubble filtering procedure on their N-body halo field, but with the slightly different barrier definition in eq. (14). The third panel was created by performing our bubble filtering procedure on the same N-body halo field, but ignoring density fluctuations outside of halos (i.e. setting $\langle\delta_{nl}(\mathbf{x_1}, z)\rangle_R = 0$), which we have verified give nearly identical bubble maps as our full procedure (so long as $\bar{x}_{HI}$ is fixed). The right-most panel was created using an approximate RT algorithm (Abel & Wandelt 2002; Sokasian et al. 2001, 2003) on the same halo field.

It is immediately obvious from Fig. 4 that all of the approximate maps (*first three panels*) reproduce the RT map (*right-most panel*) fairly well. Even the HII bubble filtering performed directly on the linear density field (*left-most panel*) performs well, which is encouraging, as that is the starting point for our semi-numerical procedure and we only improve on this scheme.

Figure 4 shows that our HII bubble filtering algorithm is an excellent approximation to RT. The similar algorithm proposed by Zahn et al. (2007) also performs well. In comparison, our algorithm produces somewhat more "bubbly" maps but appears to better capture the connectivity of HII regions. Both are an obvious improvement on directly filtering the linear density field.

Of course, in our full algorithm we identify halos from the linear density field (rather than from simulations), so our method consumes comparable processing time to the

one used to generate the leftmost panel in Figure 4, once the halos have been identified. Moreover, we are able to capture the "stochastic" component of the halo bias that causes the relatively large differences between the leftmost panel and the full RT simulation. That is, the algorithm used to generate the leftmost panel uses the large-scale linear density field to *predict* the distribution of halos (Zahn et al. 2005, 2007). In reality, the relation is not deterministic because of random fluctuations in the small-scale modes comprising each region. This leads to nearly Poisson scatter in the halo number densities (Sheth & Lemson 1999; Casas-Miranda et al. 2002) that can substantially modify the bubble size distribution whenever sources are rare, particularly early in reionization (Furlanetto et al. 2006a). By directly sampling the small-scale modes to build the halo distribution, we better recover this scatter (at least statistically, as illustrated by Fig. 2). Another way to include this scatter is by directly sampling halos from an N-body simulation (as in Zahn et al. 2007, or the second panel of Fig. 4), although that obviously requires much more computing power.

3.1. Ionization Maps

Now that we have demonstrated in turn the success of our halo and bubble filtering procedures, we present the resulting ionization maps when the two are combined. In Figure 5, we show 100 Mpc × 100 Mpc × 0.5 Mpc slices through our 200^3 ionization field at $z = 10$, 9, 8.25, 7.25 (*left to right across rows*). With the assumption of $\zeta = 15.1$, these redshifts correspond to $\bar{x}_{HI} = 0.89$, 0.74, 0.53, 0.18, respectively. As has been pointed out by Furlanetto et al. (2004c), the neutral fraction is the more relevant descriptor; bubble morphologies at a constant $\bar{x}_{HI}$ vary little with redshift (see also McQuinn et al.
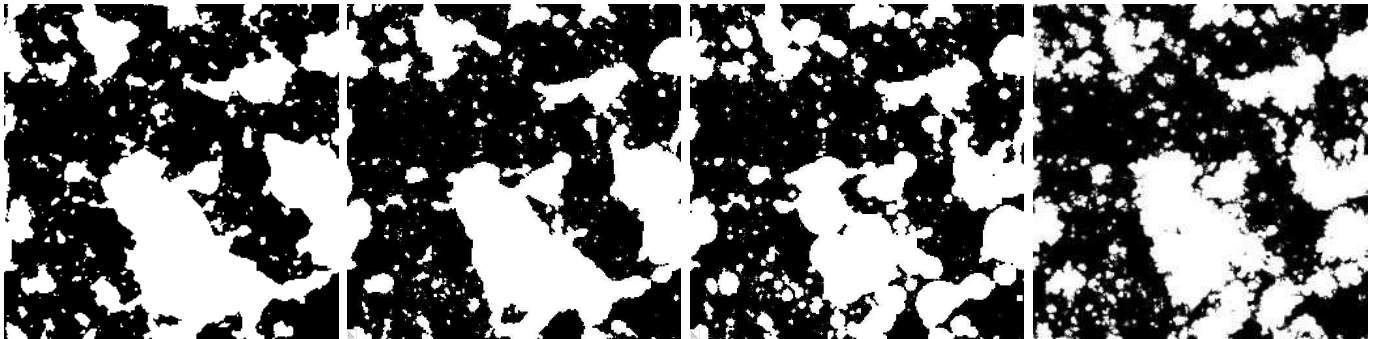
FIG. 4.— Slices from the ionization field at $z = 6.89$ created using different algorithms. All slices are 93.7 Mpc on a side and 0.37 Mpc deep, with the mean neutral fraction in the box being $\bar{x}_{HI} = 0.49$. Ionized regions are shown as white. The left-most panel was created by performing the bubble filtering procedure of Zahn et al. (2007) directly on the linear density field. The second panel was created by performing their bubble filtering procedure on their N-body halo field, but with the slightly different barrier definition in eq. (14). The third panel was created by performing our bubble filtering procedure described in § 3 on the same N-body halo field. The right-most panel (from Zahn et al. 2007) was created using an approximate RT algorithm on the same halo field.

2006a). The bottom-left panel corresponds to the halo field in the top-right panel of Fig. 3, generated on a high-resolution 1200^3 grid.

To quantify the ionization topology resulting from our method, we calculate the size distributions of both the ionized and neutral regions. We randomly choose a pixel of the desired phase (neutral or ionized), and record the distance from that pixel to a phase transition along a randomly chosen direction. We repeat this Monte Carlo procedure 10^7 times. Volume-weighted probability distribution functions (PDFs) produced thusly are shown by the solid curves in Figure 6 for ionized regions (*top panel*) and neutral regions (*bottom panel*). Curves correspond to $(z, \bar{x}_{HI}) = (10, 0.89), (9.25, 0.79), (8.50, 0.61), (8.00, 0.45), (7.50, 0.27), (7.00, 0.10)$, from left to right in the top panel, respectively (or from right to left in the bottom panel). All curves are normalized so that the probability density integrates to unity.

It is useful to compare these distributions to the analytic bubble mass function of Furlanetto et al. (2004c); although this analytic approach is motivated by the same excursion set barriers as our semi-numerical approach, it does not account for the full geometry of sources. We compute the probability distribution from the analytic model by assuming purely spherical bubbles and convolving with the volume-weighted distance to the sphere's edge:

$$p(r) \, dr = \frac{2\pi r^2 \, dr}{(1 - \bar{x}_{HI})} \int dR \, n_b(R) \left(1 - \frac{r}{2R}\right), \qquad (15)$$

where $n_b(R)$ is the comoving number density of bubbles with radii between R and $R + dR$ (taken from Furlanetto et al. 2004c).

Several points are evident from Figures 5 and 6. As expected (e.g., Furlanetto et al. 2004c, 2006a; McQuinn et al. 2006a), there is a well-defined bubble scale at each neutral fraction, despite some scatter in the sizes. This scale also gets more pronounced (i.e. the PDF peaks more) as reionization progresses; this is a result of the changing shape of the underlying matter power spectrum (Furlanetto et al. 2006a).

Also, the purely analytic estimates underpredict the size distributions at all values of the neutral fraction, though they do become increasingly accurate as the neutral fraction decreases. This trend is perhaps counterin-

tuitive, as the analytic model, which rests on the assumption of spherical bubbles, should perform best when the bubbles are isolated, as one would expect at earlier times, i.e. high neutral fractions. However, looking at the top-left panel of Fig. 5, the typical bubbles filling most of the ionized volume overlap due to the strong clustering of early sources and bubbles. This results in many "overlapping pairs of spheres" at early times, resulting from merging HII bubbles sourced by clustered sources. Thus the *spherical* bubble-based analytic model underpredicts the true size distribution, using our "mean free path" definition of bubble sizes above. This effect was not noted by previous studies (Zahn et al. 2007), because they used a different definition of bubble sizes, based on spherical filters used to flag regions in which $\bar{x}_{HI} < 0.1$. As time progresses and the universe becomes more ionized, this "overlapping pair of spheres" effect becomes less and less dominant (see Fig. 5), and the analytic model becomes increasingly more accurate.

Finally, the size distributions of neutral regions presented in the bottom panel of Fig. 6 are a new result and potentially important for the 21-cm signal (which originates in neutral hydrogen, of course). In the later stages of reionization, when the topology has transformed to isolated neutral islands in a sea of ionized gas, this figure pinpoints the typical sizes of "mostly neutral" pixels that continue to emit strongly. In contrast to the ionized regions, the neutral regions (defined in this way) do not grow substantially during reionization. From $\bar{x}_{HI} = 0.89$ to $\bar{x}_{HI} = 0.1$, the peak of the distribution shifts only by a factor of ~ 6, whereas the peak of the ionized region distribution shifts by a factor of ~ 40 over the same range. The reason for this is also evident in Figure 5: even when the universe is mostly neutral, space is dotted with islands of ionized gas, such that our "mean free path"–type size distributions never become too large. The converse does not hold true for ionized regions. However, a slight parallel for ionized islands in a mostly neutral IGM, could be found in Lyman limit systems (LLS) inside larger HII regions (e.g. Barkana & Loeb 2002; Shapiro et al. 2004; Miralda-Escudé et al. 2000), though it is not clear how prevalent such neutral clumps are at high redshifts.

Throughout this paper, we have used a $L = 100$ Mpc "simulation" box. This size facilitates comparison of our results with those from recent hybrid N-
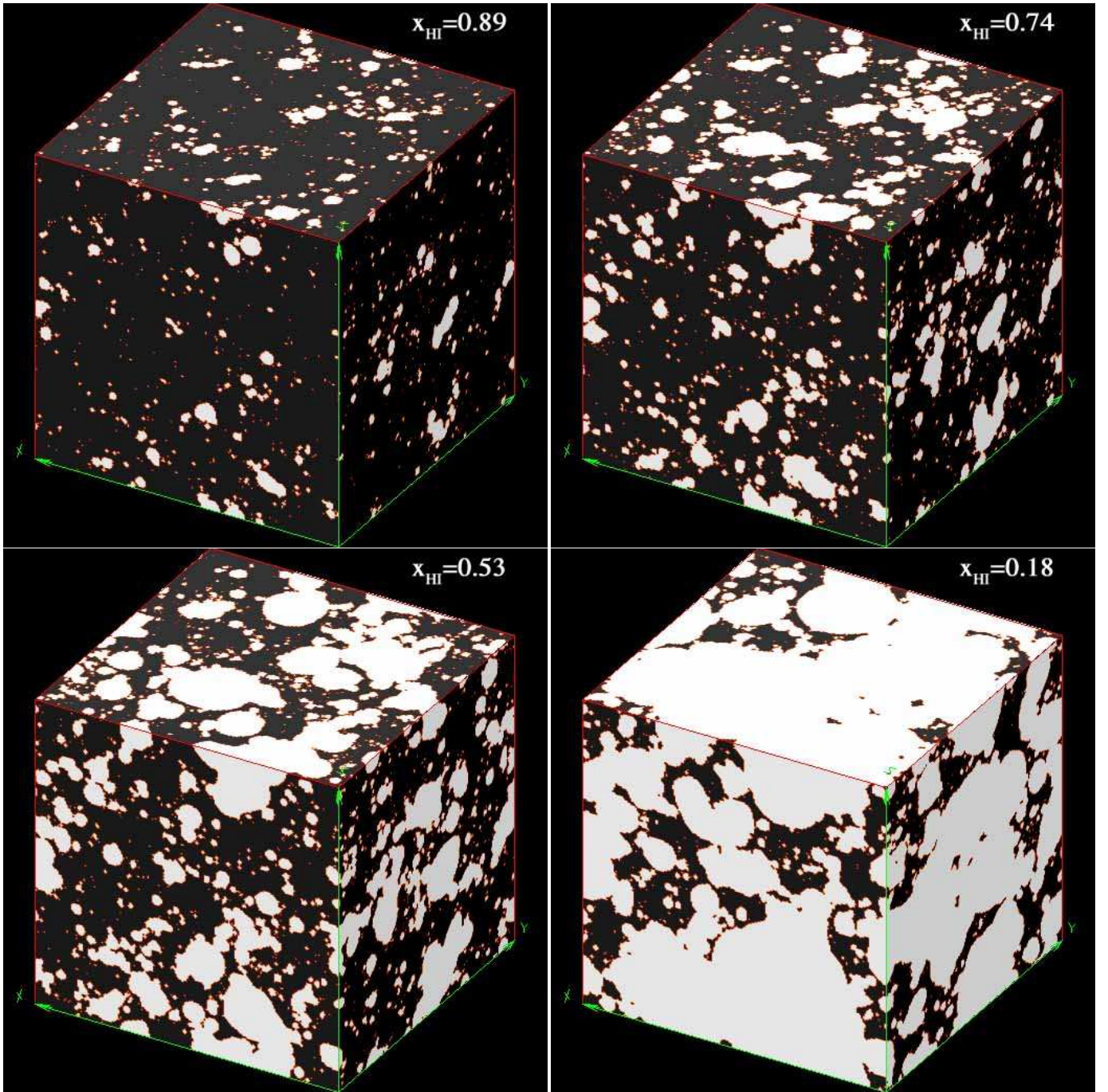
Fig. 5.— Slices through the 200^3 ionization field at $z = 10, 9, 8.25, 7.25$ (*left to right across rows*). With the assumption of $\zeta = 15.1$, these redshifts correspond to $\bar{x}_{HI} = 0.89, 0.74, 0.53, 0.18$, respectively. All slices are 100 Mpc on a side and 0.5 Mpc deep. The bottom-left panel corresponds to the halo field in the top-right panel of Fig. 3, generated on a high-resolution 1200^3 grid.

body works (Zahn et al. 2007; McQuinn et al. 2006a; Iliev et al. 2006a; Trac & Cen 2006). However, the speed of our semi-numerical approach allows us to explore larger cosmological scales while still *consistently* resolving the small halos that could dominate the photon budget during reionization. As mentioned previously, existing N-body codes must resort to merger-tree methods to populate their distribution of small-mass halos, even for box sizes $\lesssim 100$ Mpc (McQuinn et al. 2006a). In this spirit, we present some preliminary results from a $N = 1500^3$, $L = 250$ Mpc simulation, capable of directly

resolving halos with masses $M \gtrsim 2.2 \times 10^8 M_\odot$, with resulting mass functions accurate to better than a factor of two even at the smallest scale. This resolution pushes the RAM limit of our machine and so each redshift can take several hours to complete.[8]

In Figure 7, we compare size distributions of ionized

[8] We note here that our halo-finding algorithm requires significantly higher resolution than does predicting the ionization field directly from the linear density field smoothed on larger scales (Zahn et al. 2005, 2007). The latter method can be extended to even larger boxes, though at the price of a somewhat less accurate ionization map (compare the left and right panels in Fig. 4).
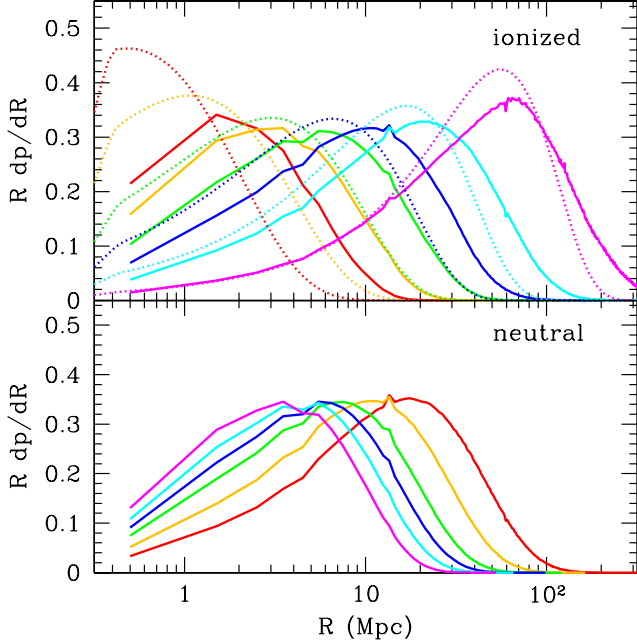
FIG. 6.— Size distributions (see definition in text) of ionized (*top panel*) and neutral (*bottom panel*) regions. Curves correspond to $(z, \bar{x}_{\rm HI}) = (10, 0.89), (9.25, 0.79), (8.50, 0.61), (8.00, 0.45), (7.50, 0.27), (7.00, 0.10)$, from left to right in the top panel, respectively (or from right to left in the bottom panel). Solid curves are produced from our simulation while dotted curves correspond to the analytic mass function. All curves are normalized so that the probability distribution integrates to unity.
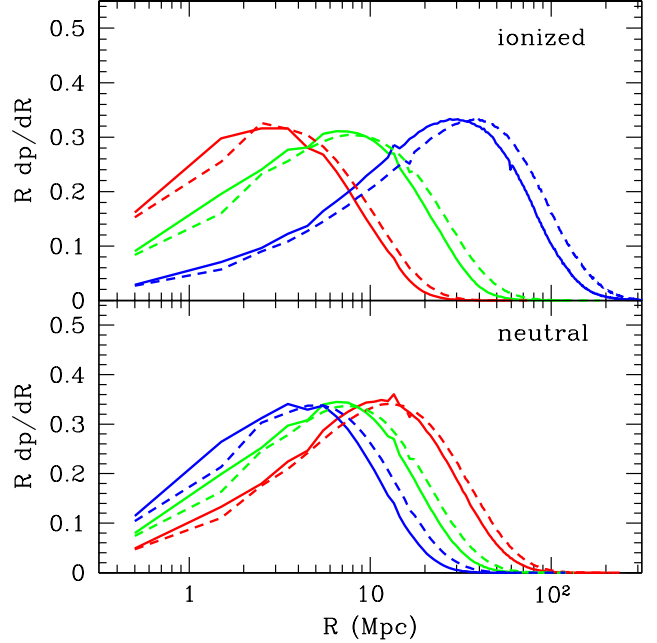
FIG. 7.— Size distributions of ionized (*top panel*) and neutral (*bottom panel*) regions from different simulation boxes. Curves correspond to $(z, \bar{x}_{\rm HI}) = (9.00, 0.80), (8.00, 0.56), (7.00, 0.21)$, from left to right in the top panel, respectively (or from right to left in the bottom panel). Solid curves are generated from our fiducial, $N = 1200^3$, $L = 100$ Mpc, simulation while dashed curves are generated from a larger simulation with $N = 1500^3$, $L = 250$ Mpc. The cell size in all ionization maps is 0.5 Mpc on a side, with the efficiency parameter, ζ, adjusted to get matching values of $\bar{x}_{\rm HI}$ and the minimum halo mass set to $M_{\min} = 2.2 \times 10^8 M_\odot$ for comparison purposes.

(*top panel*) and neutral (*bottom panel*) regions from our two different simulation boxes. Curves correspond to $(z, \bar{x}_{\rm HI}) = (9.00, 0.80), (8.00, 0.56), (7.00, 0.21)$, from left to right in the top panel, respectively (or from right to left in the bottom panel, respectively). Solid curves are generated from our fiducial, $N = 1200^3$, $L = 100$ Mpc, simulation while dashed curves are generated from our larger simulation with $N = 1500^3$, $L = 250$ Mpc. The cell size in all ionization maps is 0.5 Mpc on a side, with the efficiency parameter, ζ, adjusted to obtain matching values of $\bar{x}_{\rm HI}$, and we set the minimum halo mass to $M_{\min} = 2.2 \times 10^8 M_\odot$ even in the higher resolution runs for easier comparison.

As reionization progresses, an increasing number of large HII regions are "missed" by the $L = 100$ Mpc simulation. Interestingly, the analogous trend in the neutral region size distributions (*bottom panel*) is weaker. This is most likely because the "ionized island" effect limits the size distributions of neutral regions as described above.

4. 21-CM TEMPERATURE FLUCTUATIONS

A natural application of our "simulation" technique is to predict 21-cm brightness temperatures during reionization. The offset of the 21-cm brightness temperature from the CMB temperature, T_γ, along a line of sight (LOS) at observed frequency ν, can be written as (e.g. Furlanetto et al. 2006b):

$$\delta T_b(\nu) = \frac{T_S - T_\gamma}{1 + z}(1 - e^{-\tau_{\nu_0}}) \qquad (16)$$

$$\approx 9(1+z)^{1/2} x_{\rm HI}(1 + \delta_{\rm nl})\frac{H}{dv_r/dr + H} \quad {\rm mK},$$

where T_S is the gas spin temperature, τ_{ν_0} is the optical depth at the 21-cm frequency ν_0, $\delta_{\rm nl}$ is the physical overdensity (see discussion under eq. 14), H is the Hubble parameter, dv_r/dr is the comoving gradient of the line of sight component of the comoving velocity, and all quantities are evaluated at redshift $z = \nu_0/\nu - 1$. The final approximation makes the standard assumption that $T_S \gg T_\gamma$ for all redshifts of interest during reionization (e.g. Furlanetto 2006) and also that $dv_r/dr \ll H$. We verify in our simulation that $dv_r/dr < H$ for all neutral pixels.

Maps of $\delta T_b(\mathbf{x}, \nu)$ generated in this manner are shown in Figure 8. All slices are 100 Mpc on a side, 0.5 Mpc deep, and correspond to $(z, \bar{x}_{\rm HI}) = (9.00, 0.74), (8.25, 0.53), (7.50, 0.27)$, from left to right. The top panels take into account the velocity correction term in eq. (16), while the bottom panels ignore it.

As seen in Fig. 8, velocities typically increase the contrast in temperature maps, making hot spots hotter and cool spots cooler. We also see that temperature hot spots, which correspond to dense pixels, tend to cluster around the edges of HII bubbles, especially smaller bubbles. This occurs because HII bubbles correlate with peaks of the density field and long-wavelength biases in the density field can extend beyond the edge of the ionized region. This enhanced contrast might be useful in
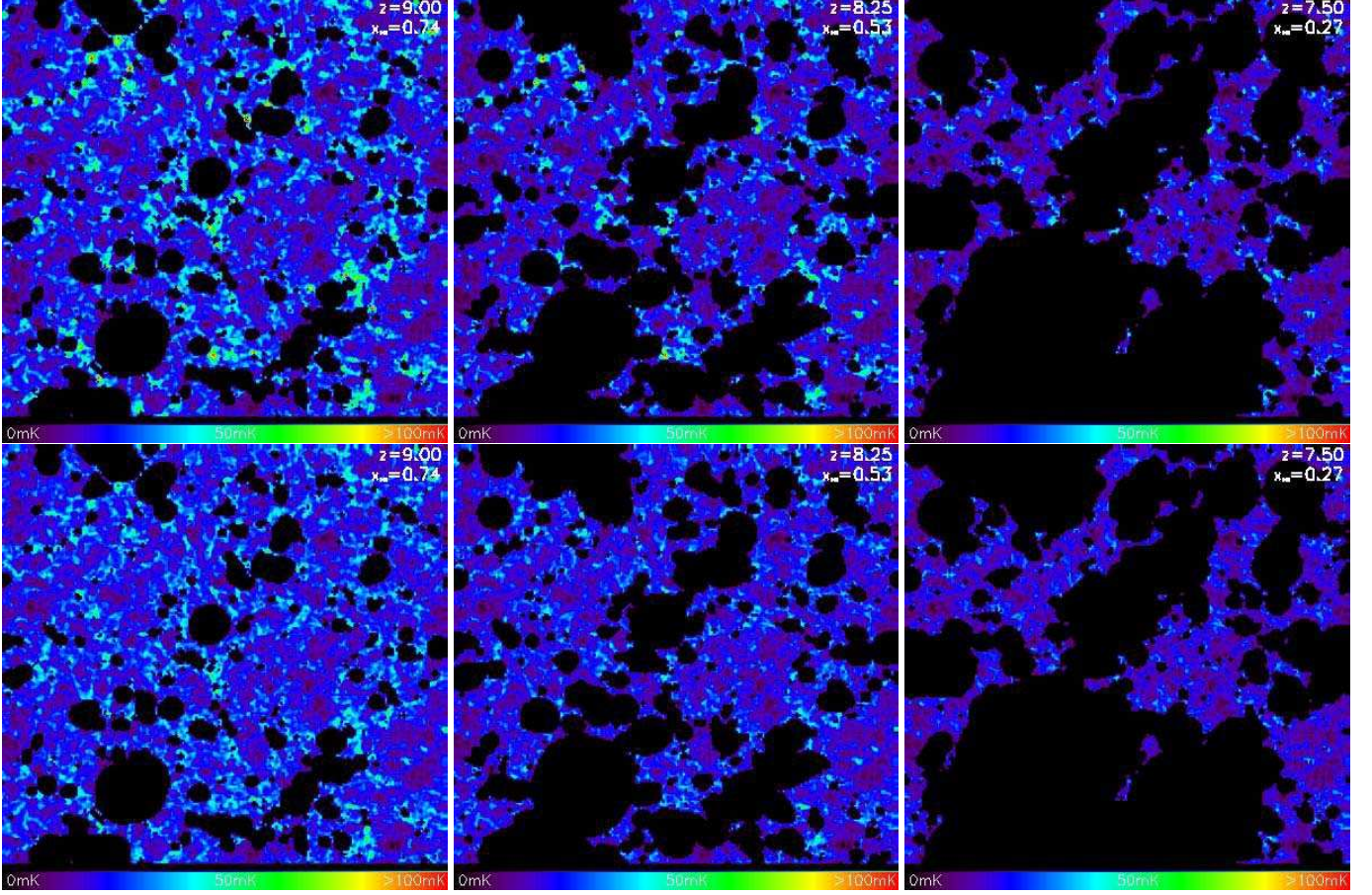
FIG. 8.— Brightness temperature of 21-cm radiation relative to the CMB temperature. All slices are 100 Mpc on a side, 0.5 Mpc deep, and correspond to $(z, \bar{x}_{\mathrm{HI}}) = (9.00, 0.74)$, $(8.25, 0.53)$, $(7.50, 0.27)$, left to right. Top panels include the velocity correction term in eq. (16), while the bottom panels do not. For animated versions of these pictures, see http://pantheon.yale.edu/~am834/Sim.

the detection of the boundaries of ionized regions with future 21-cm experiments. As reionization progresses most hot spots become swallowed up by HII bubbles, and the effects of velocities diminish.

In Figure 9 we plot the *dimensionless* 21-cm power spectrum, defined as $\Delta_{21}^2(k,z) = k^3/(2\pi^2 V) \langle |\delta_{21}(\mathbf{k},z)|^2 \rangle_k$, where $\delta_{21}(\mathbf{x},z) \equiv \delta T_b(\mathbf{x},z)/\delta\bar{T}_b(z) - 1$. Solid blue curves take into account gas velocities, while dashed red curves do not. Curves correspond to $(\bar{x}_{\mathrm{HI}}, z) = (0.79, 9.25)$, $(0.61, 8.50)$, $(0.45, 8.00)$, $(0.27, 7.50)$, $(0.10, 7.00)$, *bottom to top*. Error bars on the bottom dashed curve denote 1-σ Poisson uncertainties; fractional errors in a given bin are the same for all curves. As reionization progresses, small-scale power is traded for large-scale power, and the curves become flatter. Note that, with our dimensionless definition of the power spectrum, curves with smaller $\bar{x}_{\mathrm{HI}}$ have *larger* values of $\Delta_{21}^2(k,z)$. This is because the mean brightness temperature offset drops quite rapidly as reionization progresses, since $\delta\bar{T}_b(z) \propto \bar{x}_{\mathrm{HI}}$, but the scatter remains significant (see Fig. 6) and thus the *fractional* perturbation, $\delta_{21}(\mathbf{x},z)$, increases throughout reionization.

Finally, in Figure 10 we plot *dimensional* power spectra, $\delta\bar{T}_b(z)^2\Delta_{21}^2(k,z)$. The curves correspond to $(\bar{x}_{\mathrm{HI}}, z) = (0.80, 9.00)$, $(0.56, 8.00)$, $(0.21, 7.00)$, top to bottom at large k, respectively. The dotted green curves are generated from a large, high-resolution "simulation", with
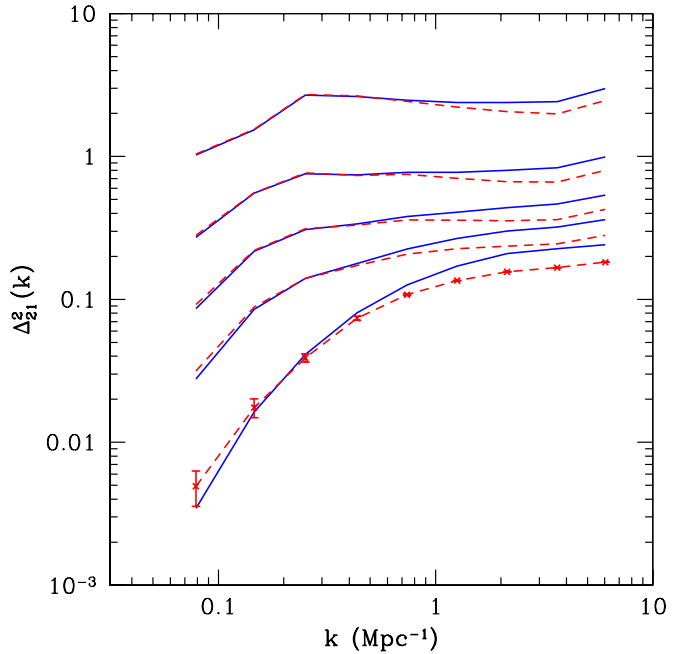

FIG. 9.— Dimensionless 21-cm power spectra for $(\bar{x}_{\mathrm{HI}}, z) = (0.79, 9.25)$, $(0.61, 8.50)$, $(0.45, 8.00)$, $(0.27, 7.50)$, $(0.10, 7.00)$, *bottom to top*. Solid blue curves take into account gas velocities, while dashed red curves do not.
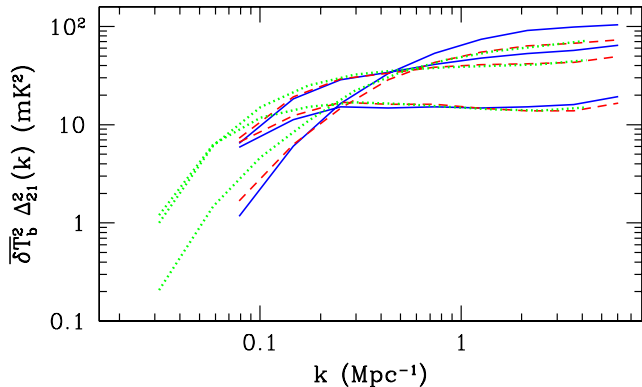
FIG. 10.— Dimensional 21-cm power spectra. The curves correspond to $(\bar{x}_{HI}, z) = (0.80, 9.00), (0.56, 8.00), (0.21, 7.00)$, top to bottom at large k, respectively. The dotted green curves are generated from a large, high-resolution "simulation", with $N = 1500^3$ and $L = 250$ Mpc, with no velocity contribution to the power spectra. Solid blue curves and dashed red curves are generated with our fiducial $N = 1200^3$ and $L = 100$ Mpc simulation, with and without the velocity contribution, respectively.

$N = 1500^3$ and $L = 250$ Mpc, with no velocity contribution to the power spectra. Solid blue curves and dashed red curves are generated with our fiducial $N = 1200^3$ and $L = 100$ Mpc simulation, with and without the velocity contribution, respectively. The cell size in all δT_b maps is 0.5 Mpc on a side, with the efficiency parameter, ζ, adjusted to achieve matching values of $\bar{x}_{HI}$ and the minimum halo mass set to $M_{\min} = 2.2 \times 10^8 M_\odot$ for comparison purposes.

As seen in Figures 9 and 10, velocities make a modest contribution to the 21 cm power spectrum, boosting power on small scales early in reionization. Note that the apparent slight decrease in power at small k when velocities are included is well within the errors from averaging over the few modes available to us on the largest scales (e.g., see Poisson error bars on the bottom dashed curve in Fig. 9). While the maximum δT_b value in our simulation box increases by a factor of a few when velocities are included, most of the pixels are only slightly affected. When the power spectrum is plotted in a dimensional version, $\delta \bar{T}_b(z)^2 \Delta_{21}^2(k, z)$, small-scale power is boosted by $\sim 40\%$ at $(\bar{x}_{HI}, z) = (0.80, 9.00)$, with this enhancement monotonically decreasing as reionization progresses. Linear theory predicts that velocities enhance the *density* power spectrum by a factor of 1.87 when $\bar{x}_{HI} = 1$ (Kaiser 1987). In fact we do recover this enhancement for a fully neutral IGM; however, as predicted by analytic models (McQuinn et al. 2006b), the ionized bubbles rapidly remove most of this amplification.

Figure 10 also confirms the inferences drawn from Fig. 7, primarily that larger box sizes are needed to capture the ionization topology at the end stages of reionization. Comparing the dashed red to the dotted green curves in Fig. 10, we note that our fiducial $L = 100$ Mpc simulations are accurate for scales smaller than $k \gtrsim 0.2$ Mpc^{-1} (or $\lambda \lesssim 30$ Mpc). As reionization progresses, larger scales lose power more rapidly than in the $L = 250$ Mpc simulation. This is again evidence that very large scale simulations are needed to model the middle and late stages of reionization. Thus the speed and high resolution of our semi-numeric approach will be extremely useful in future modeling of reionization.

5. CONCLUSIONS

We introduce a method to construct semi-numeric simulations that can efficiently generate realizations of halo distributions and ionization maps at high redshifts. Our procedure combines an excursion-set approach with first-order Lagrangian perturbation theory and operates directly on the linear density and velocity fields. As such, our algorithm can exceed the dynamic range of existing N-body codes by orders of magnitude. As this is the main limiting factor in simulating the ionized bubble topology throughout reionization, when ionized regions reach scales of tens of comoving Mpc, this will be particularly useful in such studies. Moreover, the efficiency of the algorithm will allow us to explore the large parameter space required by the many uncertainties associated with high-redshift galaxy formation.

We find that our halo finding algorithm compares well with N-body simulations on the statistical level, yielding both accurate mass functions and power spectra. We have not yet compared our halo distribution with simulations on a point-by-point basis, but we do not expect perfect agreement because of the vagaries of the excursion set approach. However, it is encouraging that a very similar algorithm independently developed by Bond & Myers (1996a) fares quite well in a comparison of high-mass halos.

Our HII bubble finding algorithm captures the bubble topology quite well, as compared to ionization maps from ray-tracing RT algorithms at an identical $\bar{x}_{HI}$. Our algorithm is similar to other codes, although we build the ionization map from our excursion set halo field rather than directly from the linear density field or from halos found in an N-body simulation (Zahn et al. 2005, 2007). Compared to codes built only from the linear density field, we can better track the "stochastic" component of the bias, though at the cost of somewhat more computation and a harder limit on resolution. On the other hand, our scheme is much faster than using an N-body code and offers superior dynamic range.

We create ionization maps using a simple efficiency parameter and compute the size distributions of ionized and neutral regions. Our size distributions are generally shifted to larger scales when compared with purely analytic models (Furlanetto et al. 2004c) at the same mean neutral fraction. The discrepancy lies in the fact that, at their core, the purely analytic models are based on ensemble-averaged distributions of isolated spheres. Hence they do not capture overlapping bubble shapes, which are most important at large $\bar{x}_{HI}$ (when the bubbles are small and random fluctuations in the source densities, as well as clustering, are most important).

In this paper, we have confined ourselves to a simple ionization criterion (essentially photon counting; Furlanetto et al. 2004c). However, our algorithm can easily accommodate more sophisticated prescriptions, so long as they can be expressed either with the excursion set formalism (Furlanetto & Oh 2005; Furlanetto et al. 2006a) or built from the halo field (in a similar way to semi-analytic models of galaxy formation embedded in numerical simulations).

We also use our procedure to generate maps and power spectra of the 21-cm brightness temperature fluctuations during reionization. We note that temperature hot spots generally cluster around HII bubbles, especially in the early phases of reionization. Because HII bubbles correlate with peaks of the density field, long-wavelength biases in the density field can extend beyond the edge of the ionized region, with the resulting overdensities appearing as hot spots. This effect might be useful for detecting the boundaries of ionized regions with future 21-cm experiments. We study the imprint of gas bulk velocities on 21-cm maps and power spectra, an effect which was not included in previous studies. We find that velocities do not have a major impact during reionization, although they do increase the contrast in temperature maps, making some hot spots hotter and some cool spots cooler. Velocities also increase small-scale power, though the effect decreases with decreasing $\bar{x}_{HI}$.

We also include some preliminary results from a simulation run with the largest dynamical range to date: a 250 Mpc box which resolves halos with masses $M \gtrsim 2.2 \times 10^8 M_\odot$. This simulation run confirms that extremely large scales are required to model the late stages of reionization, $\bar{x}_{HI} \lesssim 0.5$, when the typical scale of ionized bubbles becomes several tens of Mpc.

The speed and dynamic range provided by our semi-numeric approach will be extremely useful in the modeling of early structure formation and reionization. Our ionization maps can be efficiently folded into analyses of current and upcoming high-redshift observations, especially 21-cm surveys.

<div align="center">REFERENCES</div>

Abel, T., & Wandelt, B. D. 2002, MNRAS, 330, L53
Bagla, J. S., & Padmanabhan, T. 1997, Pramana, 49, 161
Barkana, R., & Loeb, A. 2002, ApJ, 578, 1
—. 2004, ApJ, 609, 474
Benson, A. J., Kamionkowski, M., & Hassani, S. H. 2005, MNRAS, 357, 847
Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. 1991, ApJ, 379, 440
Bond, J. R., & Myers, S. T. 1996a, ApJS, 103, 1
—. 1996b, ApJS, 103, 41
Casas-Miranda, R., Mo, H. J., Sheth, R. K., & Boerner, G. 2002, MNRAS, 333, 730
Ciardi, B., Stoehr, F., & White, S. D. M. 2003, MNRAS, 343, 1101
Dijkstra, M., Haiman, Z., Rees, M. J., & Weinberg, D. H. 2004, ApJ, 601, 666
Efstathiou, G. 1992, MNRAS, 256, 43P
Efstathiou, G., Davis, M., White, S. D. M., & Frenk, C. S. 1985, ApJS, 57, 241
Eisenstein, D. J., & Hu, W. 1999, ApJ, 511, 5
Fan, X. et al. 2006, AJ, 132, 117
Furlanetto, S. R. 2006, MNRAS, 371, 867
Furlanetto, S. R., Hernquist, L., & Zaldarriaga, M. 2004a, MNRAS, 354, 695
Furlanetto, S. R., McQuinn, M., & Hernquist, L. 2006a, MNRAS, 365, 115
Furlanetto, S. R., & Oh, S. P. 2005, MNRAS, 363, 1031
Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006b, Phys. Rep., 433, 181
Furlanetto, S. R., Zaldarriaga, M., & Hernquist, L. 2004b, ApJ, 613, 16
—. 2004c, ApJ, 613, 1
—. 2006c, MNRAS, 365, 1012
Gelb, J. M., & Bertschinger, E. 1994, ApJ, 436, 491
Gnedin, N. Y. 2000a, ApJ, 535, 530
—. 2000b, ApJ, 542, 535
Haiman, Z., & Bryan, G. L. 2006, ApJ, 650, 7
Haiman, Z., Rees, M. J., & Loeb, A. 1997, ApJ, 484, 985
Hockney, R. W., & Eastwood, J. W. 1988, Computer simulation using particles (Bristol: Hilger, 1988)
Iliev, I. T. et al. 2006a, MNRAS, submitted, preprint astro-ph/0603199
Iliev, I. T., Mellema, G., Pen, U.-L., Merz, H., Shapiro, P. R., & Alvarez, M. A. 2006b, MNRAS, 369, 1625
Jenkins, A., Frenk, C. S., White, S. D. M., Colberg, J. M., Cole, S., Evrard, A. E., Couchman, H. M. P., & Yoshida, N. 2001, MNRAS, 321, 372
Kaiser, N. 1987, MNRAS, 227, 1
Kashikawa, N. et al. 2006, ApJ, 648, 7
Kitayama, T., & Ikeuchi, S. 2000, ApJ, 529, 615
Lacey, C., & Cole, S. 1993, MNRAS, 262, 627
Liddle, A. R., Lyth, D. H., Viana, P. T. P., & White, M. 1996, MNRAS, 282, 281
Malhotra, S., & Rhoads, J. E. 2004, ApJ, 617, L5
—. 2006, ApJ, 647, L95
McQuinn, M., Lidz, A., Zahn, O., Dutta, S., Hernquist, L., & Zaldarriaga, M. 2006a, ArXiv Astrophysics e-prints
McQuinn, M., Zahn, O., Zaldarriaga, M., Hernquist, L., & Furlanetto, S. R. 2006b, ApJ, 653, 815
Mesinger, A., & Haiman, Z. 2004, ApJ, 611, L69
—. 2006, ArXiv Astrophysics e-prints astro-ph/0610258
Mesinger, A., Perna, R., & Haiman, Z. 2005, ApJ, 623, 1
Miralda-Escudé, J., Haehnelt, M., & Rees, M. J. 2000, ApJ, 530, 1
Mo, H. J., & White, S. D. M. 1996, MNRAS, 282, 347
Page, L. et al. 2006, ArXiv Astrophysics e-prints astro-ph/0603450
Peebles, P. J. E. 1980, The large-scale structure of the universe (Research supported by the National Science Foundation. Princeton, N.J., Princeton University Press, 1980. 435 p.)
Press, W. H., & Schechter, P. 1974, ApJ, 187, 425
Razoumov, A. O., Norman, M. L., Abel, T., & Scott, D. 2002, ApJ, 572, 695
Reed, D., Gardner, J., Quinn, T., Stadel, J., Fardal, M., Lake, G., & Governato, F. 2003, MNRAS, 346, 565
Scannapieco, E., Ferrara, A., & Madau, P. 2002, ApJ, 574, 590
Scoccimarro, R., & Sheth, R. K. 2002, MNRAS, 329, 629
Shapiro, P. R., Giroux, M. L., & Babul, A. 1994, ApJ, 427, 25
Shapiro, P. R., Iliev, I. T., & Raga, A. C. 2004, MNRAS, 348, 753
Sheth, R. K., & Lemson, G. 1999, MNRAS, 304, 767
Sheth, R. K., Mo, H. J., & Tormen, G. 2001, MNRAS, 323, 1
Sheth, R. K., & Pitman, J. 1997, MNRAS, 289, 66
Sheth, R. K., & Tormen, G. 1999, MNRAS, 308, 119
Sirko, E. 2005, ApJ, 634, 728
Sokasian, A., Abel, T., Hernquist, L., & Springel, V. 2003, MNRAS, 344, 607
Sokasian, A., Abel, T., & Hernquist, L. E. 2001, New Astronomy, 6, 359
Spergel, D. N. et al. 2006, ApJ, submitted, preprint astro-ph/0603449
Springel, V., & Hernquist, L. 2003, MNRAS, 339, 312
Thoul, A. A., & Weinberg, D. H. 1996, ApJ, 465, 608
Totani, T., Kawai, N., Kosugi, G., Aoki, K., Yamada, T., Iye, M., Ohta, K., & Hattori, T. 2006, PASJ, 58, 485
Trac, H., & Cen, R. 2006, ArXiv Astrophysics e-prints astro-ph/0612406
Wyithe, J. S. B., & Loeb, A. 2004, Nature, 427, 815

14

Zahn, O., Lidz, A., McQuinn, M., Dutta, S., Hernquist, L.,
    Zaldarriaga, M., & Furlanetto, S. R. 2007, ApJ, 654, 12
Zahn, O., Zaldarriaga, M., Hernquist, L., & McQuinn, M. 2005,
    ApJ, 630, 657

Zel'Dovich, Y. B. 1970, A&A, 5, 84